



PRODUCT COMPONENT

DEFINITION	
<i>Name</i>	Google Search Appliance (GSA)
<i>Description</i>	Google Search Appliance (GSA) is a combination of hardware and software in one self-contained computer system. A simple process of plugging into both a network and an electrical outlet is all that is required to setup the physical equipment.
<i>Rationale</i>	As the state moves more of its services available through the Internet, it is imperative that the State have a fully functioning and quick responding search engine. This product will provide specific abilities to track and report on activities of the Web sites being visited. The State can focus on trends of use by citizens and visitors.
<i>Benefits</i>	<ul style="list-style-type: none"> GSA is licensed for a specific quantity of indexed documents for a two year period. Our current contract is for 500,000 documents. The quantity of indexed documents is expandable (for additional fees.) It comes as an independent hardware appliance ready to plug-in to an existing network. The software is a search engine (Googlebot – a Web crawler) which explores Web sites through hyperlinks and indexes Web pages and their content. GSA produces 508 compliant Web search and results pages. GSA provides searches available by keyword, Boolean statements (searches with the “and”, “or”, “not” connectors or combination), alternate spellings for misspelled words, indexes full documents, allows restricted searches by language, allows Boolean searches, offers the ability to search within existing search results, offers the ability to customize excludes, reflects the speed of the search results, has the ability to search by alphanumeric characters, and respects “robot.txt” files which allow/disallow permissions for directories/documents within Web sites. Searches can include results from specific document types such as Microsoft Word documents, Adobe Acrobat PDF (portable file documents), and 218 other file types. Administration of the GSA can be managed by a central administration team (Office of Administration – Information Technology Services Division). Google allows the customization of specific Web sites or sections of Web sites into a collection; and the layout (known as front_ends) of the Web site to be incorporated into the GSA’s search and results pages. State agency Webmasters can manage their collections, front_ends and statistics. Search across secure content using Google’s Authorization – API to integrate into existing access control systems. Citizens have an appreciation for the added-value of this GSA being implemented and used within the State of Missouri’s Web pages. Their level of satisfaction has increased significantly.
ASSOCIATED ARCHITECTURE LEVELS	
<i>Specify the Domain Name</i>	Interface
<i>Specify the Discipline Name</i>	Branding
<i>Specify the name of the associated Technology Area</i>	Search Engines

KEYWORDS			
<i>List Keywords</i>	Search, search engine, ranking, Google, indexing, find, collection, collections, front_end, front_ends, Boolean, keyword, keywords, links, GSA, crawl, Googlebot, Internet, Intranet, search appliance, query, queries, phrases, alternate spellings, customization		
VENDOR INFORMATION			
<i>Vendor Name</i>	Onix Networking Corp.	<i>Website</i>	www.onixnet.com
<i>Contact Information</i>	Tim Hunt, Onix Networking Corp., 26931 Detroit Rd, Westlake OH 44145		
POTENTIAL COMPLIANCE SOURCES			
<i>Name</i>	Google, Inc.	<i>Website</i>	www.google.com
<i>Contact Information</i>	See vendor information above.		
<i>Name</i>		<i>Website</i>	
<i>Contact Information</i>			
COMPONENT REVIEW			
<i>List Desirable aspects</i>	<ul style="list-style-type: none"> • Flexible searches, fast results, well structured results, customizable front_ends and collections. • GSA is one of the industry's leading search engines. • GSA is easily expanded by purchasing additional license increments. • New content is crawled on an on-going basis. • GSA is 508 compliant (RSMo 191.863). 		
<i>List Undesirable aspects</i>	<ul style="list-style-type: none"> • Fee based limits maximum number of documents that can be indexed at 500,000 currently. The next license is for 1.0 million documents. • The limitation of documents indexed causes its use to be restricted to the Internet and not the Intranet. (Solution: Increase the license to cover 1.0 million or more documents thus allowing state agencies to utilize GSA in place of other purchased products, i.e. InfoSeek, etc.) • Redundancy of backup systems is limited until primary appliance is down. If the appliance fails, the backup system requires front_ends and collections to be loaded, and then the Googlebot crawls the Web to re-collect items and index them. A process period of approximately 24-hours is required before the search engine is fully functional again. 		
<i>Operating System</i>	Proprietary	<i>Platform</i>	Standalone appliance
ASSOCIATED COMPLIANCE COMPONENTS			
Product			
<i>List the Product-specific Compliance Component Names</i>			
Configuration Links			
<i>List the Configuration-specific Compliance Component Names</i>			
COMPONENT CLASSIFICATION			
<i>Provide the Classification</i>	<input type="checkbox"/> <i>Emerging</i> <input checked="" type="checkbox"/> <i>Current</i> <input type="checkbox"/> <i>Twilight</i> <input type="checkbox"/> <i>Sunset</i>		

<i>Sunset Date</i>			
COMPONENT SUB-CLASSIFICATION			
Sub-Classification	Date	Additional Sub-Classification Information	
<input type="checkbox"/> <i>Technology Watch</i>			
<input type="checkbox"/> <i>Variance</i>			
<input type="checkbox"/> <i>Conditional Use</i>			
RATIONALE FOR COMPONENT CLASSIFICATION			
<i>Document the Rationale for Component Classification</i>			
MIGRATION STRATEGY			
<i>Document the Migration Strategy</i>			
IMPACT POSITION STATEMENT			
<i>Document the Position Statement on Impact</i>	The State of Missouri realized a huge benefit from the implementation of this search tool capable of returning satisfactory, quality results. This reliable search tool yielded useful links with relevance and date priorities. Customizable collections allows for the narrowing of a search focus.		
AGENCIES			
<i>List the Agencies Currently Utilizing this Product</i>	Office of Administration		
CURRENT STATUS			
<i>Provide the Current Status</i>	<input type="checkbox"/> <i>In Development</i> <input type="checkbox"/> <i>Under Review</i> <input checked="" type="checkbox"/> <i>Approved</i> <input type="checkbox"/> <i>Rejected</i>		
AUDIT TRAIL			
<i>Creation Date</i>	02/06/06	<i>Date Approved / Rejected</i>	3/14/06
<i>Reason for Rejection</i>			
<i>Last Date Reviewed</i>		<i>Last Date Updated</i>	
<i>Reason for Update</i>			